Math 545 - Assignment 2

Emir Sevinc - 260682995

October 12, 2018

2.2

We have $E[X_t] = cos(\omega t)E[A] + sin(\omega t)E[B] = 0 + 0 = 0$. Does not depend on t. And $\gamma_x(t+h,t) = E[X_{t+h}.X_t] - E[X_{t+h}]E[X_t] = E[X_{t+h}.X_t]$ since $E[X_t] = 0$.
$E[X_{t+h}.X_t] =$
$E[(Acos(\omega(t+h)) + Bsin(\omega(t+h))) * (Acos(\omega(t)) + Bsin(\omega(t)))]$
$= E[A^2]cos(\omega(t+h))cos(\omega(t)) + E[AB]cos(\omega(t+h))sin(\omega(t)) + E[AB]sin(\omega(t+h))cos(\omega(t)) + E[B^2]sin(\omega(t+h))sin(\omega(t))$
We have that $E[A^2] = Var(A) + E[A]^2 = Var(A) = 1 = E[B^2]$, and
$E[AB] = E[A]E[B] = 0$ as A and B are uncorrelated. So the expression reduces to $cos(\omega(t+h))cos(\omega(t)) + sin(\omega(t+h))sin(\omega(t))$. But since $cos(x)cos(y) + sin(x)sin(y) = cos(x-y)$, this equals $cos(\omega(t+h) - \omega t) = cos(\omega t + \omega h - \omega t) = cos(\omega h)$. This does not depend on t either, thus we conclude that X_t is stationary. Since $cos(\omega h)$ is the autocovariance function of a stationary time series, it must be non-negative definite.

2.9

a.

$E[Y_t] = E[X_t] + E[W_t] = 0 + 0 = 0$ since we know that AR(1) process has 0 mean and W_t is white noise. For autocovariance, we have that $Y_t = X_t + W_t$. Since $E[Y_t] = 0$, $Cov[Y_{t+h}, Y_t] = E[Y_{t+h}.Y_t]$
$= E[(X_{t+h} + W_{t+h})(X_t + W_t)] = E(X_{t+h}.X_t) + E(X_{t+h}.W_t) + E(W_{t+h}.X_t) + E(W_{t+h}.W_t)$. We have shown in class that AR(1) is a linear process, thus a linear function of Z_t, and since Z is uncorrelated with W, X is also uncorrelated with W. We also have that X and W have 0 means, so we have $E(X_{t+h}.X_t) + E(X_{t+h})E(W_t) + E(W_{t+h})E(X_t) + E(W_{t+h}.W_t) = E(X_{t+h}.X_t) + 0 + 0 + E(W_{t+h}.W_t) = E[X_{t+h}.X_t] + E[W_{t+h}.W_t]$

If h is not 0, final expression is equal to $\gamma_x(h) + \gamma_w(h)$. Since W is white nosie, at nonzero h its autocovariance equals 0, and we have shown in class that the autocovariance function for an AR(1) process, that is $\gamma_x(h)$ with Z as the white noise component is equal to $\frac{\phi^h \sigma_z^2}{1-\phi^2}$. Thus the ACVF at $h \neq 0$ is $\frac{\phi^h \sigma_z^2}{1-\phi^2}$
.

If $h = 0$, then still none of the previously canceling terms remain, but now the expression we get equals $E[X_t.X_t] + E[W_t.W_t] = Var(X_t) + Var(W_t)$. We know from class that $\gamma_x(0) =$ Variance of an AR(1) process is with Z as its WN component is $\frac{\sigma_z^2}{1-\phi^2}$. Thus the expression simplifies to $\frac{\sigma_z^2}{1-\phi^2} + \sigma_w^2$. None of these have t dependence, so we have that Y is starionary.

b.

Letting $U_t = Y_t - \phi Y_{t-1}$. We need to find $\gamma_u(h)$ for $|h| > 1$.Since Y_t has 0 mean, $E[U_t] = E(Y_t) + \phi E[Y_{t-1}] = 0$ and thus $\gamma_u(h) = E[U_{t+h}, U_t]$. This is equal to $E[(Y_{t+h} - \phi Y_{t-1+h})(Y_t - \phi Y_{t-1})]$
$= E[Y_{t+h}.Y_t] - \phi E[Y_{t+h}.Y_{t-1}] - \phi E[Y_{t+h-1}.Y_t] + \phi^2 E[Y_{t+h-1}.Y_{t-1}]$
$= \gamma_y(h) - \phi\gamma_y(h+1) - \phi\gamma_y(h-1) + \phi^2\gamma_y(h)$.

For $|h| > 1$ Plugging in our $\gamma_y(h)$ value at $h \neq 0$ gives $\frac{\phi^h \sigma_z^2}{1-\phi^2} - \frac{\phi.\phi^{h+1}\sigma_z^2}{1-\phi^2} - \frac{\phi.\phi^{h-1}\sigma_z^2}{1-\phi^2} + \frac{\phi^2.\phi^h\sigma_z^2}{1-\phi^2}$
$= \frac{\phi^h \sigma_z^2}{1-\phi^2} - \frac{\phi^{h+2}\sigma_z^2}{1-\phi^2} - \frac{\phi^h \sigma_z^2}{1-\phi^2} + \frac{\phi^{h+2}\sigma_z^2}{1-\phi^2} = 0$. So by definition, this is 1-correlated and thus MA(1).

2.13

a

It is shown in example 2.4.4 that for AR(1), $w_{ii} = (1-\phi^{2i})(1+\phi^2)(1-\phi^2)^{-1} - 2i\phi^{2i}$. $\hat{\rho}(1)$ was approximated to be 0.438, Thus a 95% confidence interval for $\rho(1)$ is given by $0.438 \pm \frac{1.96[(1-\phi^2)(1+\phi^2)(1-\phi^2)^{-1}-2\phi^2]^{1/2}}{10}$.

SImilarly one for $\rho(2)$ is given by $0.145 \pm \frac{1.96[(1-\phi^4)(1+\phi^2)(1-\phi^2)^{-1}-4\phi^4]^{1/2}}{10}$.

We know that the correlation funciton of an AR(1) process is given by $\rho(h) = \phi^h$. Thus $\rho(1) = \phi$. If the model has $\phi = 0.8$, we have that $\rho(1) = 0.8$ as well. Let's see if this fits in our confidence interval. Plugging $\phi = 0.8$ simplifies our confidence interval to (0.3204,0.556). This does not contain our known value of $\rho(1)$, that is 0.8, so not a good fit. We also have $\rho(2) = \phi^2 = 0.8^2 = 0.64$. Once again letting the parameter of the model $\phi = 0.8$, the confidence interval for $\rho(2)$ simplifies to (-0.55,0.345). The actual value of 0.64 is not caught by the interval. Both of these suggest that the data may not be compatible with an AR(1) model of this parameter.

b

It was shown in class that for an MA(1) model, $w_{ii} = 1 - 3\rho^2(1) + 4\rho^4(1)$ if i = 1, and $1 + 2\rho^2(1)$ if i > 1. So for $\rho(1)$ the confidence interval is $0.438 \pm \frac{1.96[1-3\rho^2(1)+4\rho^4(1)]^{1/2}}{10}$. Similarly for $\rho(2)$, the confidence interval is constructed as $0.145 \pm \frac{1.96[1+2\rho^2(1)]^{1/2}}{10}$. For an MA(1) process we have shown in class that $\rho(h) = \frac{\theta}{1+\theta^2}$ for $h = \pm 1$. Thus $\rho(1) = \frac{0.6}{1.36} = 0.4411$. Plugging this value into the first CI gives (0.308,0.5677). The real value 0.4411 is within this bound. We also have that for an MA(1) process, $\rho(2) = 0$ as h is greater than 1. The interval numercally becomes (-0.11,0.402) by substituting $\rho(1) = 0.4411$ into the second CI. The real value of $\rho(2) = 0$ is in this bound. This suggests that the data is indeed compatible for MA(1) with parameter 0.6.

2.15

We have that the squared error S is equal to $E[(X_{n+1} - \hat{X}_{n+1})^2]$, and we need to minimize this. Suppose that the predictor $\hat{X}_{n+1} = a_0 + a_1 X_n + ... + a_n X_1$. So now,
$S = E[(X_{n+1} - a_0 - a_1 X_n - ... - a_n X_1)^2]$. Rearranging for a better square form;
$= E[(X_{n+1} - a_1 X_n - ... - a_n X_1 - a_0)^2]$
$= E[(X_{n+1} - a_1 X_n - ... - a_n X_1)^2] - 2a_0 E[X_{n+1} - a_1 X_n - ... - a_n X_1] + a_0^2$
$= E[(X_{n+1} - a_1 X_n - ... - a_n X_1)^2] + a_0^2$. To minimize, we set the derivatives equal to 0.

$\frac{\partial S}{\partial a_0} = 2a_0$
$\frac{\partial S}{\partial a_1} = -2E[X_n(X_{n+1} - a_1 X_n - ... - a_n X_1)]$
$\frac{\partial S}{\partial a_2} = -2E[X_{n-1}(X_{n+1} - a_1 X_n - ... - a_n X_1)]$
.
.
.
$\frac{\partial S}{\partial a_i} = -2E[X_{n+1-i}(X_{n+1} - a_1 X_n - ... - a_n X_1)]$

Setting this to 0, gives $a_0 = 0$, and $E[(X_{n+1} - \hat{X}_{n+1}).X_{n+1-i}] = 0$

The question says $X_t = \phi_1 X_{t-1} + ... + \phi_p X_{t-p} + Z_t$. Letting $t = n+1$ we get $X_{n+1} = \phi_1 X_n + ... + \phi_p X_{n+1-p} + Z_{n+1}$. Substituting this for X_{n+1}, and the previously assumed $\hat{X}_{n+1}$ in the above expression gives $E[(\phi_1 X_n + ... + \phi_p X_{n+1-p} + Z_{n+1} - a_1 X_n - ... - a_n X_1).X_{n+1-i}] = 0$. $\hat{X}_{n+1}$ has n terms, and X_{n+1} has p terms. Thus, if we let $a_i = \phi_i$ for $i \leq p$, and 0 else, this will hold. The resulting expression is precisely the predictor provided, thus it must be the best one. The MSE is equal to $E[(X_{n+1} - \hat{X}_{n+1})^2] = E[Z_{n+1}^2] = \sigma^2$.

2.21

We know that for an MA(1) model such as this, we have the following:
$\gamma_x(h) = \sigma^2(1+\theta^2)$ if $h = 0$
$\gamma_x(h) = \sigma^2\theta$ if $h \pm 1$
$\gamma_x(h) = 0$ if $|h| > 1$

And $E[X_t] = 0$ for all t. These will be used throughout.

a.

To predict X_3 in terms of X_1 and X_2, we will need to establish the following:

$W = [X_2, X_1]^T$

$$\Gamma = \begin{matrix} Cov(X_2, X_2) & Cov(X_1, X_2) \\ Cov(X_2, X_1) & Cov(X_1, X_1) \end{matrix} = \begin{matrix} \sigma^2(1+\theta^2) & \sigma^2\theta \\ \sigma^2\theta & \sigma^2(1+\theta^2) \end{matrix}$$

$\tilde{\gamma} = [Cov(X_3, X_2), Cov(X_3, X_1)]^T = \gamma_x(h) = [\sigma^2\theta, 0]^T$.

We have that $\tilde{a}$ needs to sovle $\Gamma\tilde{a} = \tilde{\gamma}$. Which means $\tilde{a} = \Gamma^{-1}\tilde{\gamma}$. We will simplify things by taking out a σ^2 from everything.

We have $det(\Gamma) = \frac{1}{(\theta^2+1)^2-\theta^2} = \frac{1}{(\theta^4+\theta^2+1)}$.

Thus $\Gamma^{-1} =$

$$\frac{1}{(\theta^4+\theta^2+1)} \begin{matrix} (1+\theta^2) & -\theta \\ -\theta & (1+\theta^2) \end{matrix}$$

$\implies \tilde{a} = \Gamma^{-1}\tilde{\gamma} = \frac{1}{(\theta^4+\theta^2+1)}[(\theta+\theta^3), (-\theta^2)]^T$.

Thus the best predictor is given by $\frac{1}{(\theta^4+\theta^2+1)}[(\theta+\theta^3)X_2 - \theta^2 X_1]$

So $a_1 = (\theta+\theta^3)\frac{1}{(\theta^4+\theta^2+1)}$ and $a_2 = (-\theta^2)\frac{1}{(\theta^4+\theta^2+1)}$. We also had $\tilde{\gamma} = [\sigma^2\theta, 0]^T$

Thus $MSE = Var(X_3) - \tilde{a}\tilde{\gamma} = \sigma^2(1+\theta^2) - \frac{\theta+\theta^3}{\theta^4+\theta^2+1}\sigma^2\theta$

b.

To predict X_3 in terms of X_4 and X_5, we will need to establish the following:

$W = [X_5, X_4]^T$

$$\Gamma = \begin{matrix} Cov(X_5, X_5) & Cov(X_4, X_5) \\ Cov(X_5, X_4) & Cov(X_4, X_4) \end{matrix} = \begin{matrix} \sigma^2(1+\theta^2) & \sigma^2\theta \\ \sigma^2\theta & \sigma^2(1+\theta^2) \end{matrix}$$

$\tilde{\gamma} = [Cov(X_3, X_5), Cov(X_3, X_4)]^T = \gamma_x(h) = [0, \sigma^2\theta]^T$.

We have that $\tilde{a}$ needs to sovle $\Gamma\tilde{a} = \tilde{\gamma}$. Which means $\tilde{a} = \Gamma^{-1}\tilde{\gamma}$. We will simplify things by taking out a σ^2 from everything.

We have $det(\Gamma) = \frac{1}{(\theta^2+1)^2-\theta^2} = \frac{1}{(\theta^4+\theta^2+1)}$.

Thus $\Gamma^{-1} =$

$$\frac{1}{(\theta^4+\theta^2+1)}\begin{matrix}(1+\theta^2) & -\theta \\ -\theta & (1+\theta^2)\end{matrix}$$

$\implies \tilde{a} = \Gamma^{-1}\tilde{\gamma} = \frac{1}{(\theta^4+\theta^2+1)}[((-\theta^2),(\theta+\theta^3)]^T$.

Thus the best predictor is given by $\frac{1}{(\theta^4+\theta^2+1)}[(\theta+\theta^3)X_4 - \theta^2 X_5]$

So $a_2 = (\theta+\theta^3)\frac{1}{(\theta^4+\theta^2+1)}$ and $a_1 = (-\theta^2)\frac{1}{(\theta^4+\theta^2+1)}$. We also had $\tilde{\gamma} = [0, \sigma^2\theta]^T$

Thus $MSE = Var(X_3) - \tilde{a}\tilde{\gamma} = \sigma^2(1+\theta^2) - \frac{\theta+\theta^3}{\theta^4+\theta^2+1}\sigma^2\theta$

c.

To predict X_3 in terms of X_1, X_2, X_4 and X_5 we will need to establish the following:

$W = [X_2, X_1]^T$

$$\Gamma = \begin{matrix} Cov(X_5,X_5) & Cov(X_4,X_5) & Cov(X_2,X_5) & Cov(X_1,X_5) \\ Cov(X_5,X_4) & Cov(X_4,X_4) & Cov(X_2,X_4) & Cov(X_1,X_4) \\ Cov(X_5,X_2) & Cov(X_4,X_2) & Cov(X_2,X_2) & Cov(X_1,X_2) \\ Cov(X_5,X_1) & Cov(X_4,X_1) & Cov(X_2,X_1) & Cov(X_1,X_1) \end{matrix} = \begin{matrix} \sigma^2(1+\theta^2) & \sigma^2\theta & 0 & 0 \\ \sigma^2\theta & \sigma^2(1+\theta^2) & 0 & 0 \\ 0 & 0 & \sigma^2(1+\theta^2) & \sigma^2\theta \\ 0 & 0 & \sigma^2\theta & \sigma^2(1+\theta^2) \end{matrix}$$

$\tilde{\gamma} = [Cov(X_3,X_5), Cov(X_3,X_4), Cov(X_3,X_2), Cov(X_3,X_1)]^T = \gamma_x(h) = [0, \sigma^2\theta, \sigma^2\theta, 0]^T$.

We have that $\tilde{a}$ needs to sovle $\Gamma\tilde{a} = \tilde{\gamma}$. Which means $\tilde{a} = \Gamma^{-1}\tilde{\gamma}$. We will simplify things by taking out a σ^2 from everything.

$$\Gamma^{-1} = \begin{matrix} \frac{(1+\theta^2)}{(1+\theta^2)^2-\theta^2} & \frac{\theta}{\theta^2-(1+\theta^2)^2} & 0 & 0 \\ \frac{\theta}{\theta^2-(1+\theta^2)^2} & \frac{(1+\theta^2)}{(1+\theta^2)^2-\theta^2} & 0 & 0 \\ 0 & 0 & \frac{(1+\theta^2)}{(1+\theta^2)^2-\theta^2} & \frac{\theta}{\theta^2-(1+\theta^2)^2} \\ 0 & 0 & \frac{\theta}{\theta^2-(1+\theta^2)^2} & \frac{(1+\theta^2)}{(1+\theta^2)^2-\theta^2} \end{matrix}$$

$\implies \tilde{a} = \Gamma^{-1}\tilde{\gamma} = [\frac{\theta^2}{\theta^2-(1+\theta^2)^2}, \frac{\theta(1+\theta^2)}{(1+\theta^2)^2-\theta^2}, \frac{\theta(1+\theta^2)}{(1+\theta^2)^2-\theta^2}, \frac{\theta^2}{\theta^2-(1+\theta^2)^2}]^T$.

Thus the best predictor is given by $\frac{\theta^2}{\theta^2-(1+\theta^2)^2}X_5 + \frac{\theta(1+\theta^2)}{(1+\theta^2)^2-\theta^2}X_4 + \frac{\theta(1+\theta^2)}{(1+\theta^2)^2-\theta^2}X_2 + \frac{\theta^2}{\theta^2-(1+\theta^2)^2}X_1$
$= \frac{1}{\theta^4+\theta^2+1}[-\theta^2 X_5 + (\theta+\theta^3)X_4 + (\theta+\theta^3)X_2 - \theta^2 X_1]$. We note the similarity of this expression to the previous cases.

Thus $MSE = Var(X_3) - \tilde{a}\tilde{\gamma} = \sigma^2(1+\theta^2) - 2\frac{\theta+\theta^3}{\theta^4+\theta^2+1}\sigma^2\theta$